Report on the 4th International Workshop on Large-Scale Graph Data Analytics (LSGDA 2025)

Zhengyi Yang
University of New South Wales
Sydney, Australia
zhengyi.yang@unsw.edu.au

Dong Wen
University of New South Wales
Sydney, Australia
dong.wen@unsw.edu.au

Wentao Li
University of Leicester
Leicester, United Kingdom
wl226@leicester.ac.uk

ABSTRACT

This report summarizes the program and outcomes of the 4th International Workshop on Large-Scale Graph Data Analytics (LSGDA 2025). The workshop was held in conjunction with the VLDB 2025 conference in London, UK, on September 5, 2025. The aim of the workshop was to provide a forum for researchers from academia and industry to exchange ideas, techniques, and application scenarios in large-scale graph data analytics, as well as to discuss open challenges and identify new research directions in the area. The program featured eleven research-track presentations and one encore-track presentation, fostering rich discussions and collaborations among participants.

1 FOREWORD

We are delighted to present the program and outcomes of the 4th International Workshop on Large-Scale Graph Data Analytics (LS-GDA), held in conjunction with the International Conference on Very Large Databases (VLDB) in London, UK on September 5, 2025.

Over the past few years, LSGDA has established itself as a key event for the exchange of innovative ideas and research findings in the area of large-scale graph data analytics. This year's workshop builds on the success of previous editions [1], held in conjunction with ICDE 2019 in Macau, VLDB 2020 in Tokyo, and VLDB 2024 in Guangzhou. The growing interest and participation in LSGDA underscore the importance and relevance of graph data analytics in various application domains such as social networks, communication networks, biological networks, and transportation networks. These domains naturally generate large-scale graph data, necessitating the development of novel and scalable analytics techniques.

The workshop aims to bring together people from around the world with different backgrounds to exchange frontier research ideas and results in large-scale graph data analytics. It also aims to integrate techniques from various areas into solving problems in graph data analytics by focusing on research topics such as machine learning techniques for graphs and AI techniques for graphs. The proposed workshop aligns well with the interests of VLDB 2025. It is closely related to several research tracks in VLDB 2025, including Graph Data Management, Social Networks, Recommendation Engines, Data Mining and Analytics, Provenance and Workflows, Spatial, Temporal, and Multimedia Databases, Scientific and Medical Data Management, and Profile-based or Context-Aware Data Management. The workshop is also expected to spark discussions on novel technique paradigms and application scenarios for graph data analytics.

We received 20 submissions this year, and through a rigorous review process by three members of the program committee for each paper, we have selected 12 papers to be presented in the workshop, including 1 encore paper and 11 research papers. The papers feature a diverse array of topics, including graph data modeling, graph database query processing, mining techniques, parallel processing, and machine learning for graphs. We believe that the research presented here will not only advance the field of large-scale graph data analytics but also inspire new ideas and future research directions. The program committee for LSGDA2025 includes senior researchers and rising stars of the data management community from both academia and industry. We are grateful to the committee members for their thorough evaluation of the submissions within the short review time.

We are honored to have Prof. M. Tamer Özsu (University of Waterloo), Prof. Christian S. Jensen (Aalborg University), A/Prof. Da Yan (Indiana University Bloomington), and invited speakers from ByteDance to deliver keynote talks at the workshop. We express our gratitude to the VLDB 2025 organizers for their support which makes LSGDA possible. We also thank all the authors, reviewers, and participants for their contributions and commitment to making this workshop a success. We are deeply appreciative of the financial support we have received from our diamond sponsor **ByteDance**.

2 ORGANIZATION

General Chairs
- Wenjie Zhang, University of New South Wales
- Ying Zhang, Zhejiang Gongshang University

Program Committee Chairs
- Zhengyi Yang, University of New South Wales
- Dong Wen, University of New South Wales
- Wentao Li, University of Leicester

Encore Chair
- Qingqiang Sun, Great Bay University

Publicity Chairs
- Hanchen Wang, University of Technology Sydney
- Michael Ruisi Yu, Chinese University of Hong Kong

Web Chair

- Mingchen Ju, University of New South Wales

Program Committee Members

- Alexander Zhou, Hong Kong Polytechnic University
- Cheng Chen, ByteDance
- Furqan Aziz, University of Leicester
- Kaisong Huang, University of Calgary
- Kongzhang Hao, Google
- Miguel Rodriguez, Google
- Raghavender Puchhakayala, JPMorgan
- Tianming Zhang, Zhejiang University Of Technology
- Xu Zhou, Hunan University
- Yifan Wang, University of Hawaii
- Yunkai Lou, Alibaba Group
- Zihan Yang, University of Southern Queensland

3 KEYNOTES

- Prof. M. Tamer Özsu (University of Waterloo)
- Prof. Christian S. Jensen (Aalborg University)
- A/Prof. Da Yan (Indiana University Bloomington)
- Industry talk from ByteDance

4 ACCEPTED PAPERS

(1) "Top-r Influential Community Search in Bipartite Graphs"
Yanxin Zhang, Zhengyu Hua, Long Yuan

(2) "Hyracks Unchained: Efficient Recursion for Navigational Queries in Apache AsterixDB"
Glenn Galvizo, Michael Carey

(3) "GAL: Topology-Aware Serialization for Graph Traversals"
Zeynep Korkmaz, Tamer Özsu, Khuzaima Daudjee

(4) "Experiment & Benchmark Paper: To What Extent Does Quality Matter? The Impact of Graph Data Quality on GNN Model Performance"
Jana Vatter, Maurice L. Rochau, Ruben Mayer, Hans-Arno Jacobsen

(5) "EnGraph: Ensemble-Based Augmentation for Graph Anomaly Detection"
Andrew Shields, Robert Sheehy, Pat Doody

(6) "Semantic Embedding for Enterprise Clustering: A Systematic and Scalable Approach Using SentenceTransformers"
Yigong Xiao, Xianzhi Lei, Kecheng Wang, Changan Zhou, Niannian Huang

(7) "Shape-Aware, Scale-Agnostic Representation of Dynamic DAGs"
Jennifer Neumann, Peter M. Fischer

(8) "Vision Paper: Improving the Accessibility of Port Operations in Supply Chain Management using Graph Data Analysis"
Mert Ayas, Frank Laarmann, Leif Meier, Katja Zeume

(9) "Growing Up HAL: Historic and Property Graph Queries"
Muhammad Khan, Ioana Manolescu, Angelos-Christos Anadiotis

(10) "Efficient Betweenness Maximization in Temporal Networks"
Xijuan Liu, Kejia Xu, Lele Zhang, Haiyang Hu, Ying Zhang

(11) "Single-Source Regular Path Querying in Terms of Linear Algebra"
Semyon Grigorev, Georgiy Belyanin, Rodion Suvorov

REFERENCES

[1] Long Yuan, Zhengyi Yang, Qingqiang Sun, and Alexander Zhou. 2024. Report on the 3rd International Workshop on Large-Scale Graph Data Analytics (LSGDA 2024). In *Proceedings of Workshops at the 50th International Conference on Very Large Data Bases, VLDB 2024, Guangzhou, China, August 26-30, 2024.* VLDB.org. https://vldb.org/workshops/2024/proceedings/LSGDA/LSGDA24.01.pdf